ELSEVIER

Journal of Volcanology and Geothermal Research 106 (2001) 211–228

Journal of volcanology and geothermal research

www.elsevier.nl/locate/jvolgeores

Identifying and discriminating seismic patterns leading flank eruptions at Mt. Etna Volcano during 1981–1996

S. Vinciguerra[a,*], V. Latora[b,c], S. Bicciato[d], R.T. Kamimura[e]

[a]*Benfield Greig Hazard Research Centre, Department of Geological Sciences, University College London, Gower Street, London WC1E 6BT, UK*

[b]*Center for Theoretical Physics, MIT, 77 Massachusetts Avenue, Cambridge, MA, USA*

[c]*Department of Physics, Harvard University, Cambridge, MA, USA*

[d]*Dipartimento di Processi Chimici dell'Ingegneria Universita' degli Studi di Padova, Università di Padova, via Marzolo 9, 35131 Padova, Italy*

[e]*Department of Chemical Engineering, MIT, 77 Massachusetts Avenue, Cambridge, MA, USA*

Received 24 January 1999; accepted 20 October 2000

Abstract

Seismicity affecting Mt. Etna Volcano (Italy) has been investigated in order to identify and discriminate seismic patterns precursory to flank eruptions. An intense period (1981–1996) of seismicity and volcanism, during which eight flank eruptions occurred has been considered. Two statistical methods are used: mean hypothesis testing and entropic decision trees. The results of the two methods are consistent and reveal a pattern of 'deep' and 'western' events, prior to the flank eruptions that can be used as a predictive tool as well as a physical modeling constraint.

Keywords: seismic patterns; flank eruptions; Mt. Etna Volcano

1. Introduction

Etna volcano is located on the eastern coast of Sicily, close to the boundary between the continental crust of the Hyblean Plateau and the Mesozoic oceanic crust of the Ionian basin (Fig. 1).

As one of the planet's few continuously active volcanoes, Etna constitutes one of the most important natural volcanic laboratories. The complex stress field affecting the volcano results from the combined effects of: (1) regional tectonics associated with the interaction between the African and Eurasian plates; (2) transient local stresses caused by magma moving into feeder dikes and, in places, gravitational stress due to the shape and dimension of the volcano edifice. The latter results in the distinct mechanical behavior of the unstable eastern slope of the volcano, which shows a seaward displacement with respect to the western slope as shown by geological, seismological and volcanological analyses (Guest et al., 1984; Kieffer, 1985; Borgia et al., 1992; Lo Giudice and Rasà, 1992; McGuire et al., 1997; Patanè et al., 1994; Montalto et al., 1996; Rasà et al., 1996). At deeper levels, magma ascends through the crust as large bodies emplaced along a number of megafaults (Armienti et al., 1989; Ferrucci et al., 1993; McGuire et al., 1997).

Due to this complex framework, analyses of time–space evolution of seismicity affecting the volcanic area as a 'whole' have provided inconsistent results.

* Corresponding author. Tel.: +44-207-6792429; fax: +44-207-3887614.

E-mail address: s.vinciguerra@ucl.ac.uk (S. Vinciguerra).

0377-0273/00/$ - see front matter
PII: S0377-0273(00)00274-2

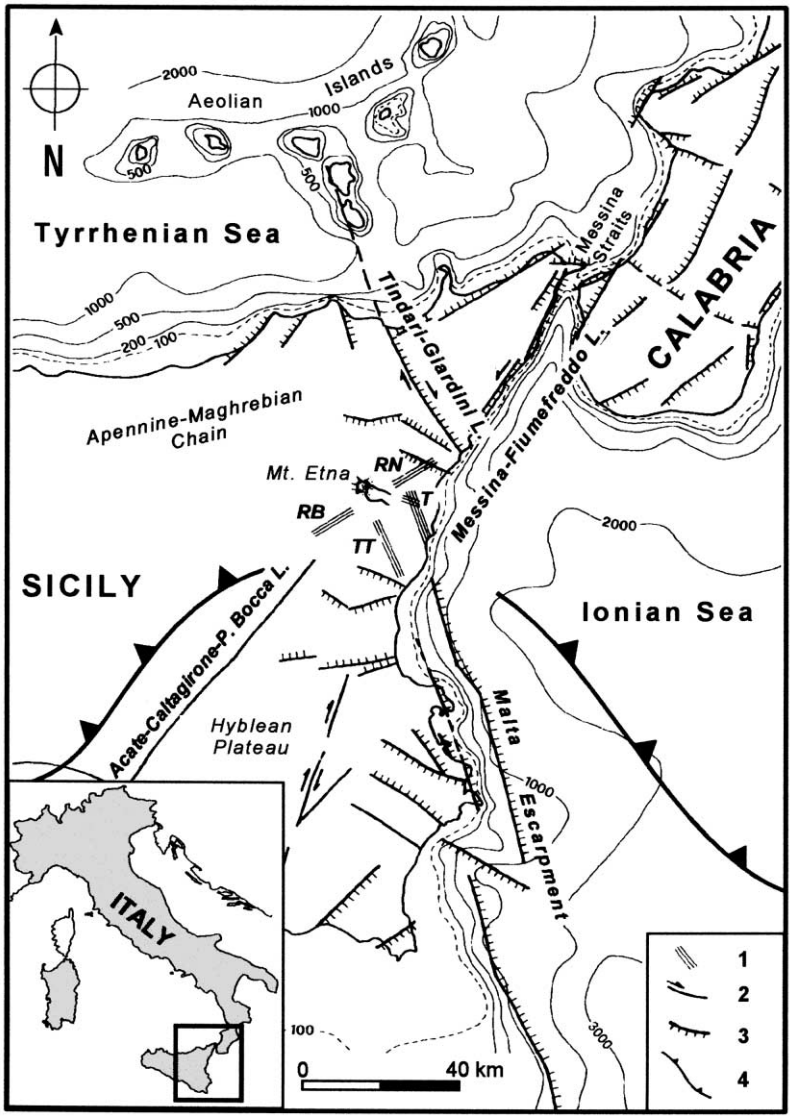

Fig. 1. Tectonic sketch map of eastern Sicily (redrawn from Lo Giudice et al., 1982); (1) main structural elements, (2) regional lineaments, (3) faults.

For example, Sharp et al. (1981), by analyzing a catalogue of historical felt earthquakes from 1532 to 1978 and by selecting 'primary' events in the Etna region, suggest that there is a statistically significant relationship between the onset of lateral eruptions and the prior occurrence of either summit eruptions or earthquakes. However, Nercessian et al. (1991) do not agree with the Sharp et al. (1981) mode of earthquakes selection, stating that the 'complete' catalogue suppresses those events considered as aftershocks and becomes more Poissonian. They conversely emphasize that the ending phase of the lateral eruption is accompanied by strong earthquakes ($M \backsim 4$), leading to redistribution of stresses and pressures in the heterogeneous mass of Mt. Etna. However, no statistically significant correlation is provided by these authors, who refer exclusively to the October 1984 strong earthquakes affecting the eastern flank. A

moderate agreement to Nercessian et al. (1991) is given by the statistical correlation found by Gresta et al. (1994) between larger Etna earthquakes ($I_o > IX$) and the end of eruptions. Nevertheless Gresta et al. (1994) state that these results agree only partially with Nercessian et al. (1991), since the largest eruptions are not followed by increases in seismicity as in October 1984. Finally Gasperini et al. (1990), by applying a generalized Poisson process of the Shlien and Toksoz (1970) type, do not find evidence for relations between the eruptive mechanisms and the occurrence of the earthquake sequences during 1978–1987.

Conversely, when defined time–space seismic patterns have been considered, evidence for links between seismicity and eruptive processes have been found. In this sense smaller shocks seem to be directly linked to the eruptive processes (Patanè et al., 1984; Cristofolini et al., 1985; Gresta and Patanè, 1987). Recent analyses, concerning cross-correlation functions applied to low-frequency events (Cardaci et al., 1993) and objective criteria in defining 'precursors' (in terms of 'false alarms', 'correct' and 'failed' predictions), have further confirmed this (Gresta and Longo, 1994). A strong link between flank eruptive history and regional state of stress has been proposed by Mulargia et al. (1991, 1992), who use a statistical pattern recognition algorithm to argue for a significant correlation between seismicity in the Gulf of Patti (off the north coast of Sicily) and the onset of flank eruptions. Moreover, time variations of the *b* coefficient of the frequency–magnitude relationship have been observed at Mt. Etna some weeks before the start of the two flank eruptions of 1981 and 1983 (Gresta and Patanè, 1983a,b). These variations have been interpreted in terms of small but rapid changes in the stress field acting on the higher parts of the volcano (Gresta and Patanè, 1983b). Furthermore, fractal analysis of earthquake time and spatial distributions (De Rubeis et al., 1997; Latora et al., 1998; Barbano et al., 1999) in different sectors of the Etna region revealed a link between variations of seismicity clustering properties at mid (order of months) and short-term (order of days) scales and eruptive processes.

Finally, recent analyses using the mean hypothesis test (MHT) procedure (Latora et al., 1999) to identify seismic patterns prior to flank eruptions suggest that a marker of the volcanic activity is given by 'deep' and 'western' flank earthquakes. These results are also supported by an alternative analysis of quantitative (fractal dimension, *b* value) seismological parameters (De Rubeis et al., 1997; Vinciguerra et al., 1999).

In this paper we present a complete study of the seismic activity affecting the volcano during the period 1981–1996. The aim of this work is to identify and discriminate any typical seismic pattern, which can be connected to flank eruptive processes. In order to do so we analyze the seismic time series using an integrated approach of two different statistical methods, the MHT and entropic decision trees (EDT). Encouraging preliminary results of the MHT (Latora et al., 1999) lead us here to test the validity of this analysis and to take the opportunity to integrate this approach with the EDT analysis, applied for the first time as a pattern recognition method in geophysics.

2. Data

The recent 1981–1996 period was characterized by intense volcanic and seismic activity. We have focused our attention on flank eruptions in an attempt to understand how seismicity preceding, accompanying, and following their occurrence differs from seismicity in periods of quiescence or strombolian activity. Eight main flank eruptions occurred during this time span. The seismic data set available consists of 5000 earthquakes recorded between January 1st, 1981, and December 31st, 1996, by permanent and temporary networks operating at Etna. This data set consists of merged data sets from the permanent networks run by the University of Catania and International Institute of Volcanology (Catania) and by temporary arrays run by the University of Cambridge (1981–1982), the University of Grenoble (1983–1984), and the Gruppo Nazionale per la Vulcanologia and Osservatorio Vesuviano (1989–1991). Seismicity was located using the HYPO71 routine (Lee and Lahr, 1975) and adopting the velocity model proposed for Etna region by Hirn et al. (1991). A test of completeness performed using the procedure of Tinti and Mulargia (1985) has shown that the data set can be considered complete for magnitudes greater than 2.5. Well constrained hypocentral locations (angular gap $< 180°$; number of stations > 5) were obtained

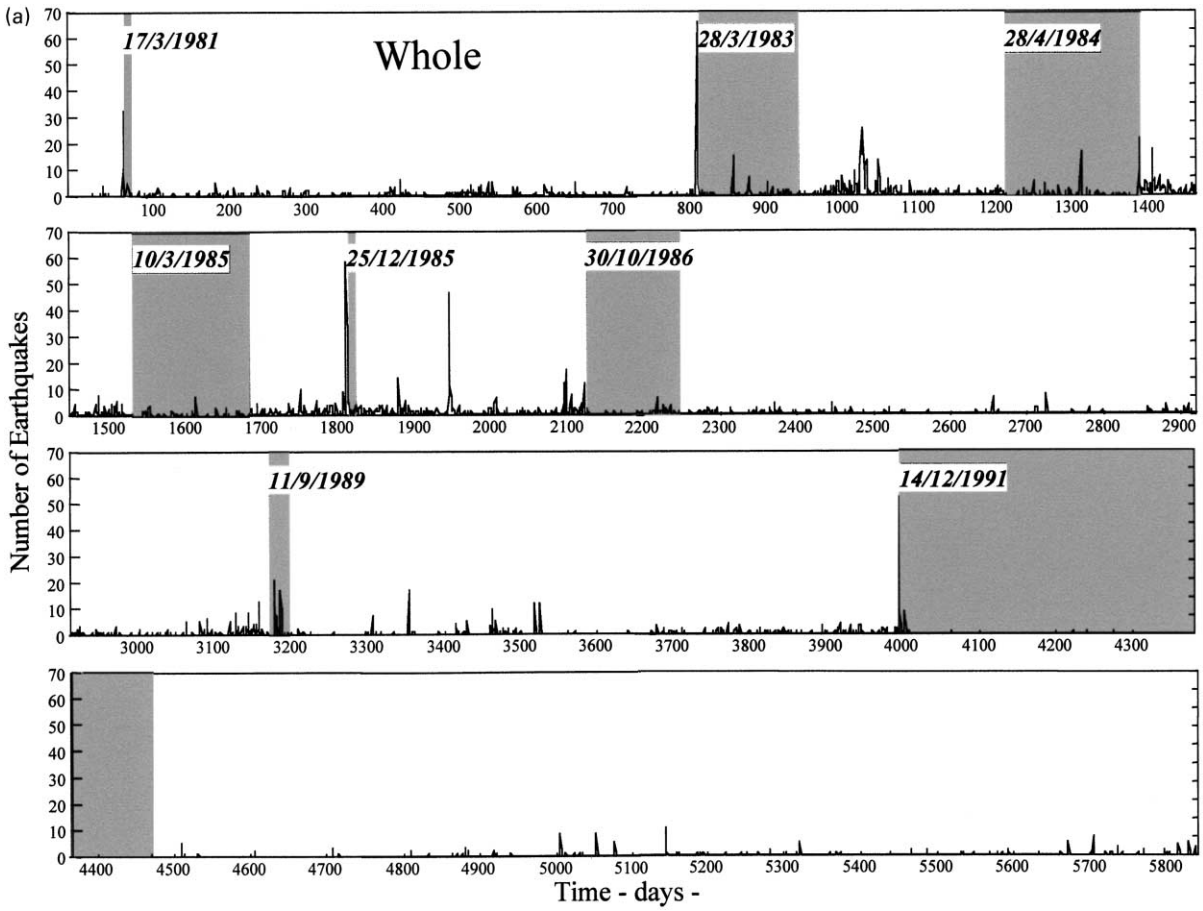

Fig. 2. Daily event counts for the *whole* (a), *deep* (b), *eastern flank* (c), *summit* (d), *western* (e) earthquake groups, during 1981–1996. The time reported on the *x*-axis is in days; day 1 corresponds to the beginning of our data set. The marked dates refer to eruption onsets; eruption duration is highlighted in gray.

for 2479 events; these were used in subsequent analyses. Magnitudes were estimated from the Serra La Nave (SLN) station seismograms by the following equation: $M_d = 2.2 \log T + 0.3 \log D - 1.5$, where M_d is a magnitude based on duration time, T is the earthquake duration in seconds and D the hypocentral distance from the SLN station in kilometres.

As a first step we studied this catalogue to define the relationships between seismicity as a 'whole' and eruptive processes. As a second step, we subdivided the total number of earthquakes into four main subgroups based on location. We did this to investigate seismic patterns affecting the main seismogenic volumes of the volcano edifice.

Seismological studies performed in the Etna region have recognized the existence of a boundary between two different crustal levels (Scarpa et al., 1983; Gresta et al., 1985; Gresta and Patanè, 1987); a shallower one ($0 < h < 7$ km) characterized by direct faulting, and a deeper level ($h > 7$ km) characterized by both direct and thrust faulting. Consequently we classified a first subgroup of events as:

(1) *Deep* earthquakes (1129 earthquakes), characterized by focal depths greater than 7 km. This depth threshold distinguishes events strictly linked to the shallow volcano dynamics from events occurring at a depth characterized by superimposed extensional regime on an older, partly relaxed and deeper compressional one (Gresta and Patanè, 1987).

The shallow events were distributed into three

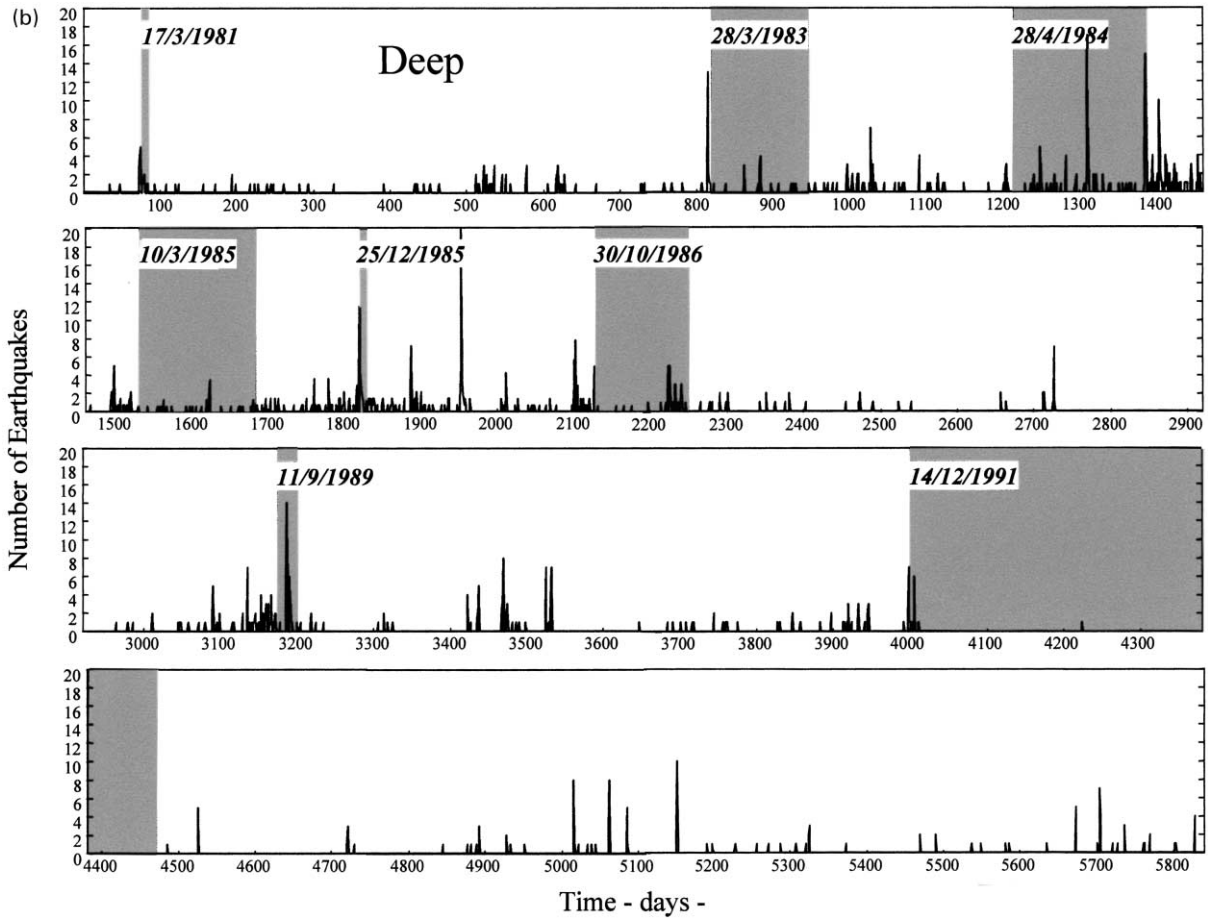

Fig. 2. (*continued*)

subgroups according to the structural framework affecting Etna:

(2) *Eastern flank* earthquakes (532 events), occurring in the unstable part of the volcano (Borgia et al., 1992; McGuire et al., 1990) and mainly associated with the NNW–SSE fault system.

(3) *Summit* earthquakes (584 events) due to rock fracturing processes induced by the penetration of magma into the shallow layers ($h > 2800$ m).

(4) *Western flank* (234 events) earthquakes related mainly to the NE–SW regional fault systems that are strongly suspected to control flank eruption occurrences (Frazzetta and Villari, 1981; Lo Giudice et al., 1982; Cristofolini et al., 1985; Bousquet et al., 1988; Gresta et al., 1990).In Fig. 2a–e we show the number of earthquakes per day between 1981 and 1996 for the total data set and the four subgroups described above.

To determine what patterns of physical phenomena distinguish eruptive from non-eruptive periods, two classes of 125-day time intervals are created:

(1) the *eruptive class* consisting of eight samples, one for each eruptive episode in the revised catalogue, with the eruption occurring at day 76;

(2) the *non-eruptive* class, created by randomly selecting eight time intervals of 125 points during non-eruptive periods.

The 125–day time window for both *eruptive* and *non-eruptive* classes was chosen based on estimates of 16–80 days required for the magma pulse to reach

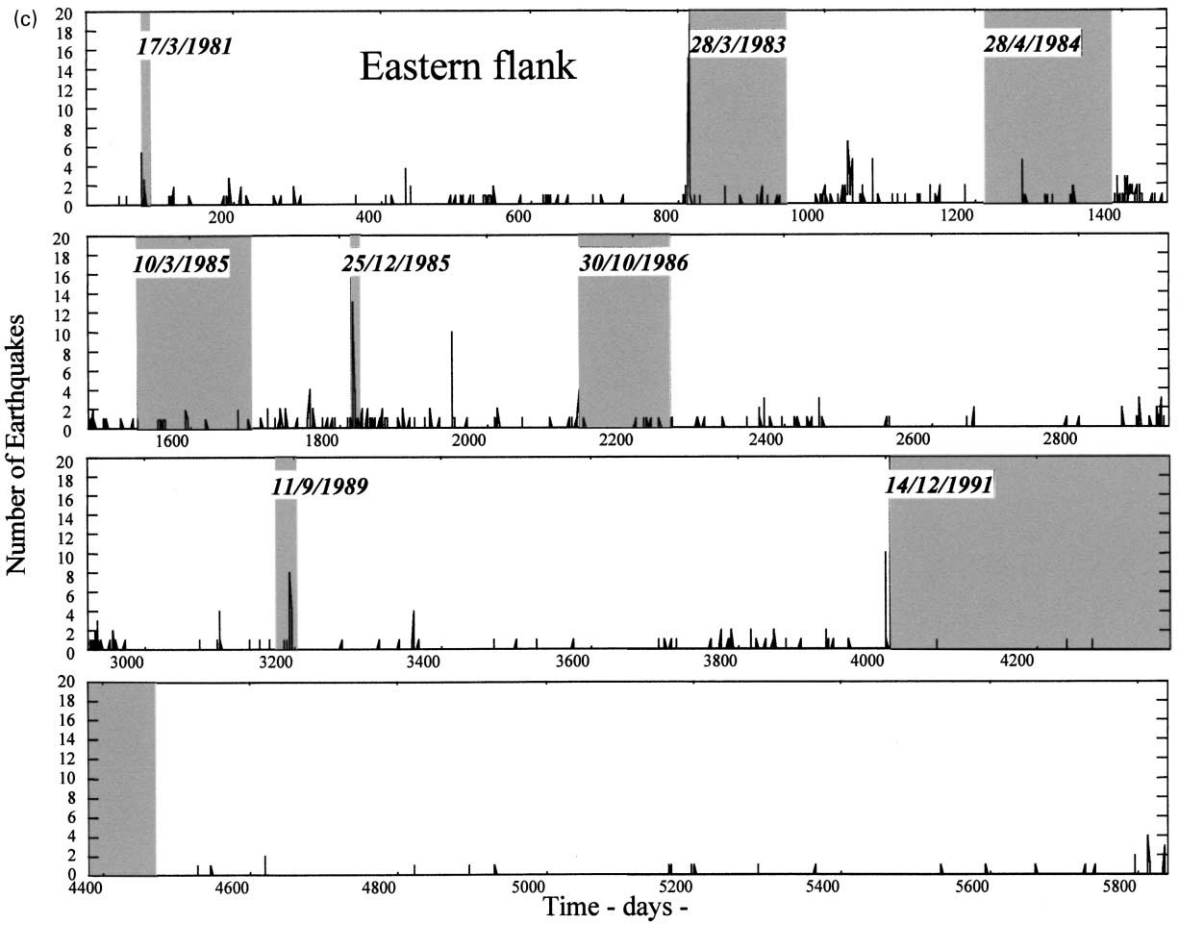

Fig. 2. (*continued*)

the surface at Kilauea (Klein et al., 1987) and at Etna (Castellano et al., 1997) before the eruption onset ($d = 76$), and on the basis of the literature, confirming the occurrence of seismic swarms occurring 2–3 months before the eruption (Gresta and Patanè, 1987; Castellano et al., 1993, 1997; Bonaccorso et al., 1996). The 49 days after the eruption onset was selected in order to guarantee the inclusion of seismic swarms linked to the opening fracture systems in the eruptive classes, since many flank eruptions during the period considered (see chronology of volcanic and seismic activity) started as sub-terminal eruptions and evolved only in a second temporal stage to a flank eruption. Moreover, the choice of 125 days for the time window length guarantees that the period analyzed should be long enough to describe intermediate time-scale events. Multiple random selections of the non-eruptive class (duplicates 1–3) have been considered to ensure a sufficient degree of generalization. Finally, the principle that intervals should not overlap, sets the maximum time length of the interval.

3. Chronology of volcanic and seismic activity

Eight main flank eruptions (March 1981, March–August 1983, April–October 1984, March–July 1985, December 1985, October 1986–March 1987, September–October 1989, December 1991–March 1993) affecting different eruptive systems (northern flank during the first eruption, southern flank during the

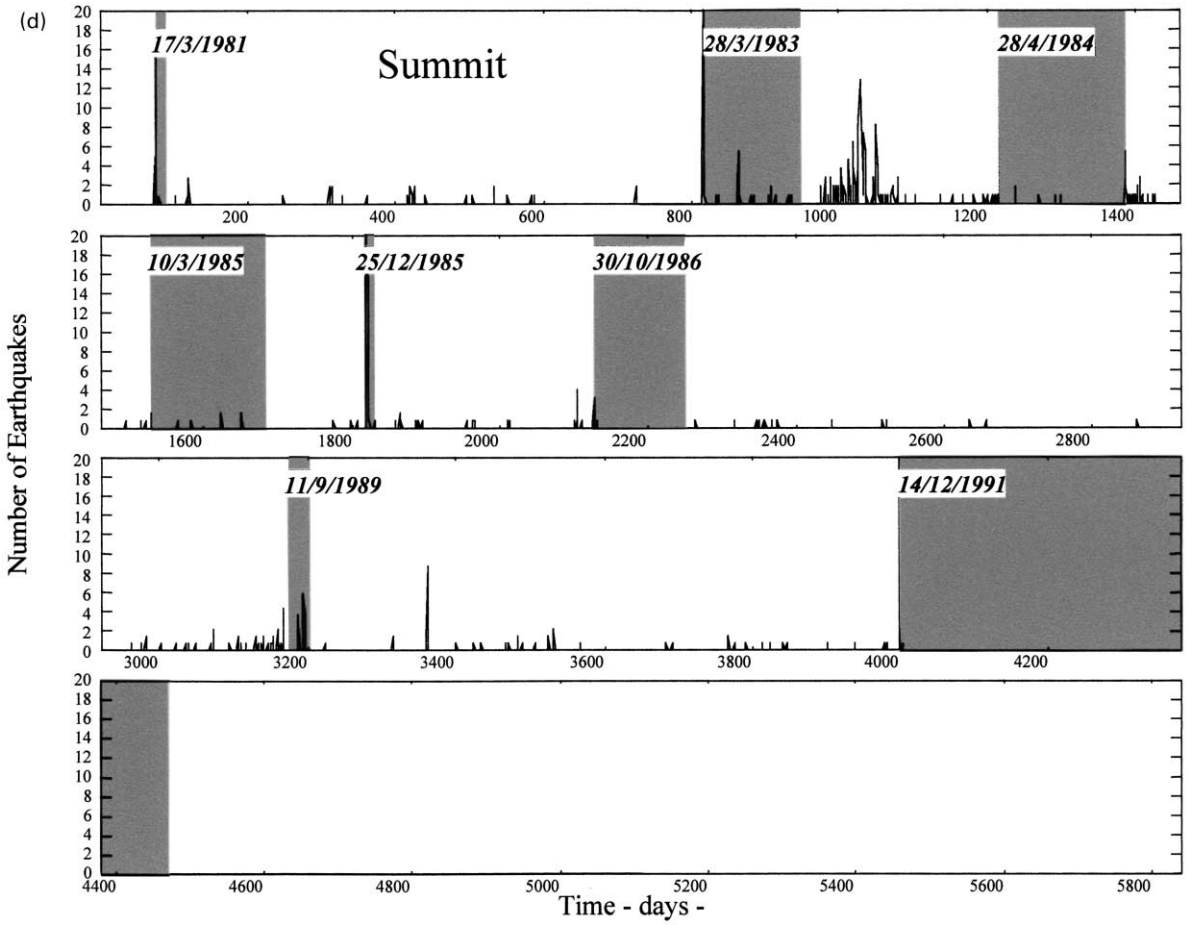

Fig. 2. (*continued*)

1983 and 1985 eruptions, and eastern flank during the others) occurred during the period considered here (Fig. 3). To better understand the geological and volcanological implications, we summarize in Table 1 the volcanic and seismic activity related to the eight main flank eruptions as reported from the Smithsonian Institute (1981, 1984, 1985, 1986, 1989, 1991) and by Azzaro and Neri (1992).

4. Method

In this section we illustrate MHT and EDT, the two statistical analysis methods used in this paper. These techniques are applied to the seismic data set and are used to identify seismic measurements that discriminate between *eruptive* and *non-eruptive* classes of behavior. The results also give important information on the relevant time periods when such discrimination is possible.

4.1. Mean hypothesis testing

MHT is a statistical test based on utilizing the mean as a measure of a class behavior (Kamimura, 1997; Latora et al., 1999). MHT seeks out those measurements that are statistically different between the two classes that are being compared. The null hypothesis, i.e. the criterion to consider if two classes are similar, is that the means are similar.

In this paper MHT has been adapted to identify which seismic measurements are more informative

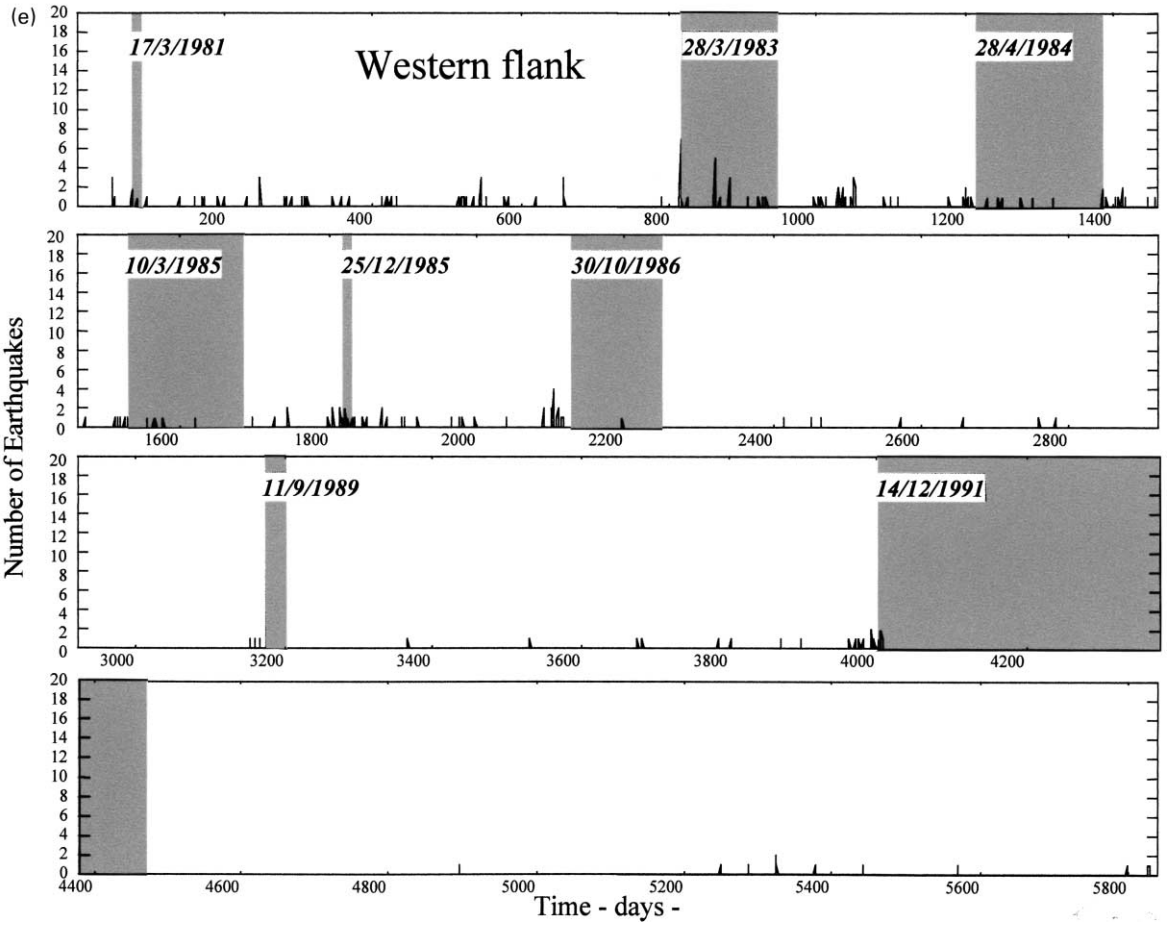

Fig. 2. (*continued*)

in discriminating between eruptive and non-eruptive volcanic periods. In order to do so, we consider two classes, *eruptive* and *non-eruptive*, as previously defined in Section 2. Each class contains eight different samples, and for each class we can study the total number of earthquakes as well as the seismic data subgroups (i.e. deep, eastern, summit, western earthquakes). By $(x_i(t))_k$, $1 \le t \le 125$ we indicate the *total* number of earthquakes per day, while by $x_{ij}(t))_k$, $1 \le t \le 125$ we indicate the number of events of the seismic data subgroup j ($j = 1, 2, 3, 4$), i.e. the number of *deep*, *eastern*, *summit* or *western* earthquakes per day, respectively. Index $i = 1$ labels the *eruptive* class and $i = 2$ labels the *non-eruptive* one; k labels the kth sample of class i, ($k = 1, \ldots, 8$). In this paper MHT is applied both to the *total* number and to each of the four subgroups separately; we explain the details of the methods in the most general case.

For each class i and for each of the seismic data subgroups j a mean is calculated at each time point averaging over the number of different samples in the following way:

$$\mu_{ij}(t) = \frac{1}{n_i} \sum_{k=1}^{n_i} (x_{ij}(t))_k \tag{1}$$

In our case we have $n_i = 8$ members (or number of samples) both for the *eruptive* class $i = 1$ and for the *non-eruptive* class $i = 2$ (see Section 2); the index k labels the kth sample of class i. In a more compact

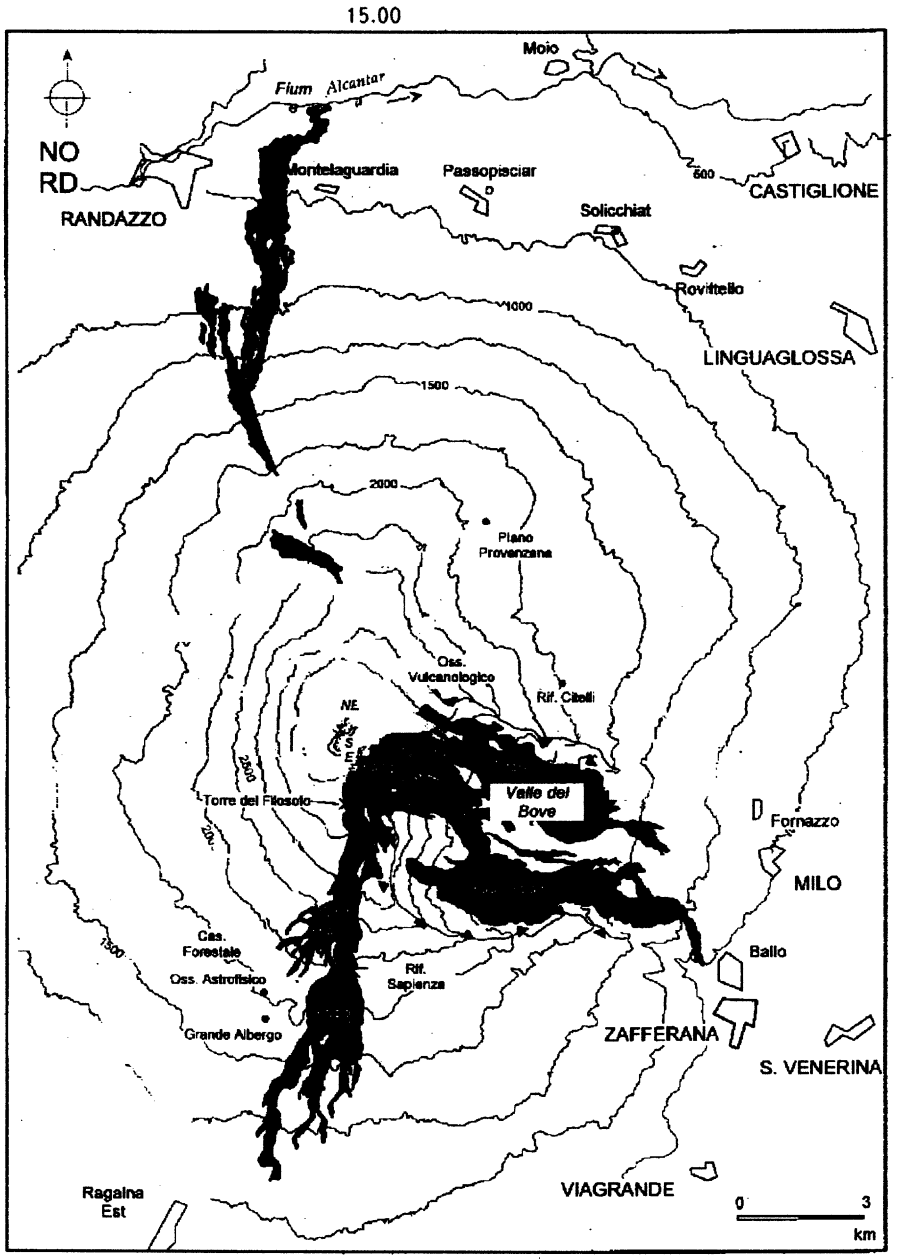

Fig. 3. Map of Mt. Etna showing extent of lava flows from 1981–1996 flank eruptions. Eruption year is marked on the correspondent lava flow. 15.00 marks the longitude used to discriminate eastern from western flanks.

way we consider the class' mean vector μ for the class i at time t which is given in terms of its components:

$$\vec{\mu}_i(t) = [\mu_{i1}(t),\ \mu_{i2}(t),\ \mu_{i3}(t),\ \mu_{i4}(t)] \qquad i = 1, 2 \quad (2)$$

Each of the components is subject to a hypothesis test at each time point to determine if the overall behavior of the *eruptive* and *non-eruptive* classes is statistically different. For a fixed t and j the null hypothesis is:

$$\mu_{1j}(t) - \mu_{2j}(t) = 0 \qquad j = 1, 2, 3, 4 \quad (3)$$

Table 1
Onset, duration, end, location and seismicity notes for the eight eruptions which occurred between 1981 and 1996

Onset	End	Duration (days)	Location	Seismicity notes
17-03-1981	23-03-1981	6	Northern flank	(1) Several hundred of earthquakes on 16 and 17 March ($2.0 < M < 3.0$) accompanied the eruption onset
28-03-1983	06-08-1983	131	Southern flank	(1) Deep earthquakes on 26 and 27 March ($2.5 < M < 3.0$) affected the western flank of the volcano, causing light damages
28-04-1984	17-10-1984	172	Summit area	(1) Deep earthquakes ($M \approx 2.5$) occurred on 25–30 September, mostly concentrated on the western flank of the volcano (2) Shallow strong earthquakes ($2.5 < M < 4.2$) on 16–19 and 25 October marked the eruption end
10-03-1985	13-07-1985	125	Summit area	(1) No increase of seismic activity occurred at the time of the eruption (2) Earthquakes recorded both before and during the eruption affected the eastern and the southern flanks
25-12-1985	31-12-1985	5	Eastern flank	(1) Swarms of events ($2.5 < M < 3.5$) at the end of October and November preceded the eruption onset (2) A strong seismic swarm ($2.5 < M < 3.0$) affecting the eastern flank, accompanied the eruption onset
30-10-1986	01-03-1987	122	Eastern flank, summit area	(1) Swarms of events ($2.5 < M < 3.3$) preceded (30 deep events with a maximum magnitude of 3.3) the eruption onset affecting the western flank (2) Other deep and shallow swarms ($2.5 < M < 3.2$) occurred on 5 and 7 October
11-09-1989	09-10-1989	29	Eastern flank, summit area	(1) Seismic activity ($2.0 < M < 3.2$) occurred during the months preceding the eruption, mainly affecting the summit area (2) In November and December seismic activity returned to low levels
14-12-1991	31-03-1993	473	Eastern flank, summit area	(1) Four swarms, affecting the upper nothern and western flanks, preceded and accompanied the first eruptive phases (2) Isolated earthquakes were recorded during 1992

the two classes' means are not statistically different and the subgroup j is not discriminating at time t.

The decision to reject or to accept the null hypothesis is made using a T-statistic (Mardia et al., 1979). The two means are used to construct the T-statistic, which is then compared to a tabulated (threshold) T-value. For simplicity, in this paper the ratio of the T-statistic to the threshold T-value, called the T-ratio, is used.

Using a selected significance level of 95%, the null hypothesis is rejected or accepted according to whether or not the T-ratio exceeds 1 (details of the entire hypothesis test can be found in Mardia et al. (1979) and Kamimura (1997). The T-ratio is a measure of how large the difference in the means are, and reflects the power of the seismic data in discriminating among *eruptive* and *non-eruptive* periods. If the T-ratio > 1, then the null hypothesis is

rejected and the differences in the means are significant. This indicates that for the variables (subgroups) selected and the time t considered, the data behave dissimilarly in the two classes. The variables are then viewed as class discriminators, or indicators of changes in the system status.

4.2. Entropic decision tree

EDT (Cheung and Stephanopoulos, 1990; Quinlan, 1986) is a statistical method to directly determine which variable(s) and variable attributes are the most discriminating in terms of their ability to segregate different class behaviors; it is based on the concept of information content of a distribution. At each node of the tree, the populations being considered are separated into two groups; to identify the conditions that make the node, i.e. the selected binary bifurcation among all the possible ones, the information content of every possible bifurcation is calculated.

The variable that maximizes the change in the information content resulting from that particular split is selected as the node for the branch point. In other words, we select the variable and the conditions that generate the best split, i.e. the highest order–lowest entropy split. This information content is determined by calculating the *information entropy* via Shannon's formula (Shannon, 1949) before and after each node candidate. The procedure is then repeated on each subsequent tree branch for the remaining variables and samples.

The way this analysis works is clearer from its concrete application to our data. We consider the same two populations examined in MHT, *eruptive* and *non-eruptive* time series $(x_1(t))_k$, $(x_2(t))_k$, where $1 \le t \le 125$, and $k = 1, 2, \ldots, 8$. Let us describe here, for simplicity, the case of the *total* number of earthquakes; exactly the same analysis will also be applied to each subgroup separately, for each of the four seismic characteristics. We have $N = n_1 + n_2 = 16$ different samples (eight corresponding to *eruptive* periods and eight to *non-eruptive* ones). Our goal is to find which day gives the best discrimination in terms of *eruptive* and *non-eruptive* classes of behavior, i.e. which day distinguishes the most eruptive samples from non-eruptive ones. From the set of all available values present in the data (in our case we have 1,2,3,… earthquakes per day) we systematically start with the first value as *discriminating value*, x_{dv}, and we observe the information entropy generated using it.

Given x_{dv}, a binary split of the N samples into two groups or branches G_1 and G_2 is generated, the first branch containing the samples with $x_i(t)_k$ 'less than' the discriminating value x_{dv}, and the second branch containing the remaining samples. Defining n_{G_l} as the total number of samples in G_l, $(l = 1, 2)$, n_{1,G_l} as the number of *eruptive* samples that fall into G_l, and n_{2,G_l} as the number of *non-eruptive* samples that fall into G_l, then the total information content after the split, $E(G_1,G_2)$, is given as the weighted sum of the information entropy of each single branch:

$$E(G_1, G_2) = \frac{n_{G_1}}{N} I(n_{1,G_1}, n_{2,G_1}) + \frac{n_{G_2}}{N} I(n_{1,G_2}, n_{2,G_2}) \tag{4}$$

where the information entropy I for each branch is defined through Shannon's formula:

$$I(n_1, n_2) = -\sum_{i=1}^{2} \frac{n_i}{N} \log\left(\frac{n_i}{N}\right). \tag{5}$$

The discriminating value and the time t producing the largest drop in information content (hence, less entropy in the branches) are selected as the most effective node. Repeating the above procedure, on the remaining times and variables, creates further branches of the tree.

We begin considering the first day $(t = 1)$ and choose 1 as the arbitrary node value x_{dv}. G_1 will contain all the samples with $x(t = 1) < 1$, i.e. all the samples for which less than 1 earthquake occurred on day 1, and G_2 will contain all other samples. We calculate the information content for this possible bifurcation. We repeat the same procedure for $x_{\mathrm{dv}} = 2, 3, \ldots$, and we select, from the information content, the best bifurcation, i.e. the best possible discrimination (in terms of *eruptive* and *non-eruptive* classes of behavior) permitted by the seismic data of day 1. The same procedure is repeated for $t = 2, 3, \ldots$ to select the day giving the best bifurcation. The ideal bifurcation (with the maximum information content) would split the 16 samples into two groups G_1 and G_2 exactly coincident with the *eruptive* and *non-eruptive* classes, respectively. In general, this is not always reached in

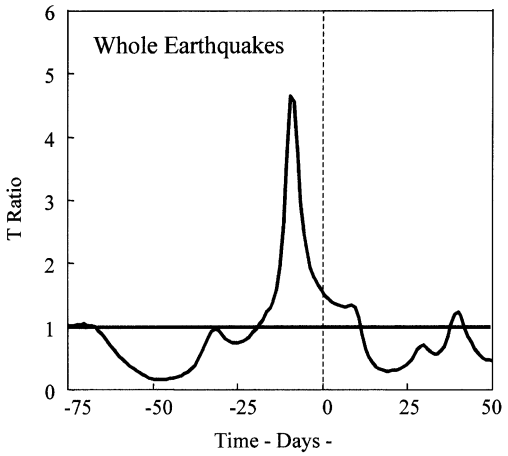

Fig. 4. Temporal evolution of the *T*-ratio for the total number of earthquakes. The dashed line indicates the onset of eruptive processes. The negative values on the time axes indicate days preceding the eruption (day 0).

one step and we need the cumulative information of more days; in this way the branching process develops.

5. Results and discussions

The results of the statistical analysis performed in this paper give us general information about the correlation between seismic activity and flank eruptions at Mt. Etna. The seismic patterns of different volcano sectors provide a basis for investigating the mechanics of the different seismogenic volumes.

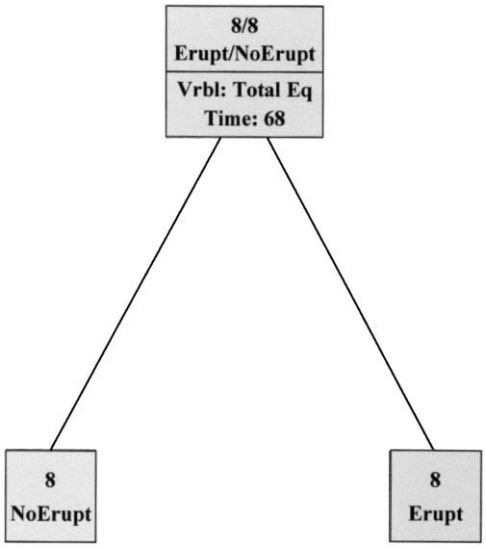

Fig. 5. Entropic decision tree for the *total* number of earthquakes.

(1) The *total* number of earthquakes gives a *T*-ratio > 1 20 days prior to eruption onset (see Fig. 4), with the highest value (around 5) found seven days prior to eruption onset. Fig. 5 shows the results of the EDT analysis for the *total* number of earthquakes. EDT sharply discriminates the *eruptive* classes from the *non-eruptive* classes at $d = 67 - 68$ (≈ 7 days prior to an eruption). The data set allows a perfect discrimination of *eruptive* and *non-eruptive* samples, based only on the information of a single day. We report in Fig. 6 the distribution of $E(G_1,G_2)$ on x_{dv}–t plane, for the best discriminating value, in order to help understand how sharp discrimination is.

When we use the other two duplicates the same sharp splitting into *eruptive* and *non-eruptive* samples is found, and the best discrimination is obtained for day $t = 67$ and $t = 68$, respectively. This means that the information contained in the seismic data series about ten days prior to an eruption can be used to perfectly discriminate eruptive from non-eruptive seismic behavior. These results indicate a clear connection between eruptive processes and seismic time series and statistically confirm the increase observed in some indicators of short-term seismicity before some of the eruptions considered by this study. Moreover these results are a further confirmation of the sharp short-term variations (order of days) of the fractal dimension found to occur before flank eruptions in the same period (Barbano et al., 1999) and of the *b* values observed three weeks before the 1981, 1983, and 1985 eruptions (Gresta and Patanè, 1983a,b; Patanè et al., 1992). These results allow us to constrain the time increase in event numbers before a flank eruption and to fix the time length (seven days) of the short-term seismic precursor of flank eruption occurrences during the 1981–1996 period.

(2) *T*-ratios for the *deep*, *summit*, *eastern*, and *western* subgroups are plotted separately in Fig. 7. Our results show that *eastern* earthquakes are not influenced by eruptive processes, while the other three subgroups are, even if in slightly different ways. In Fig. 8 we repeat the same analysis done in Fig.7, this time comparing the *eruptive* class to a

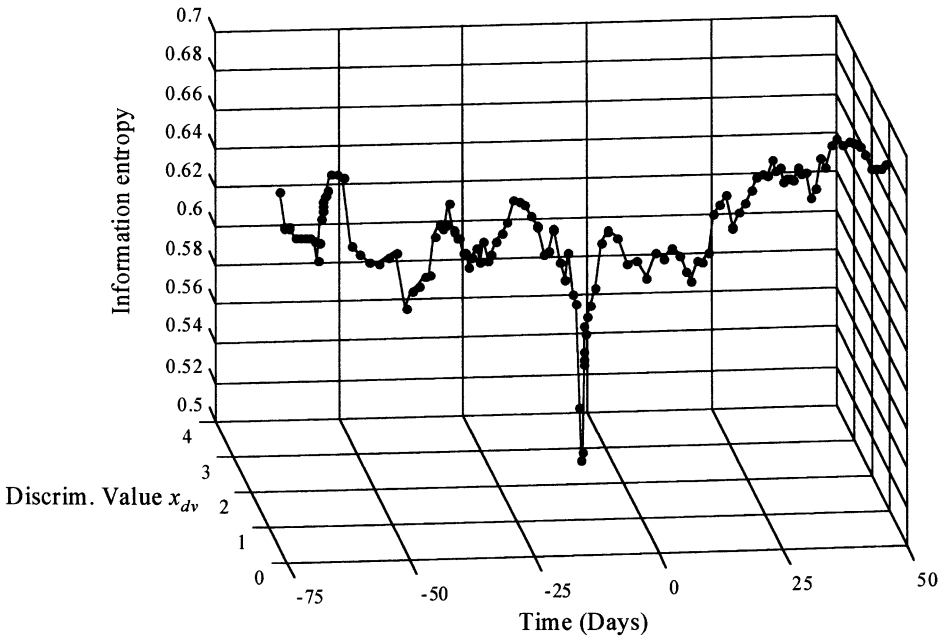

Fig. 6. Distribution of $E(G_1,G_2)$ on x_{dv}–t plane, for the best discriminating value.

different set of eight samples for the *non-eruptive* class (number 2 of the 3 duplicates of Section 2). The fact that Fig. 7 and Fig. 8 show the same behaviors is a confirmation of the robustness of our results, which do not depend on the choice of the elements of the *non-eruptive* class; this is also true for duplicate 3, which is not shown in any figure.

In Fig. 9 we report the results of the EDT analysis for the four subgroups *deep*, *summit*, *eastern* and *western*. Each subgroup allows a discrimination based on the seismic behavior before the onset of the eruptive processes $t = 65$ for deep, $t = 70$ for summit, $t = 75$ for east, and $t = 55$ for west. The discrimination is marked for all except the eastern subgroup. In detail:

(2a) For the *deep* earthquakes the *T*-ratio has a value greater than 1 for the 20 to 0 days before the eruption onset. EDT discriminates the *eruptive* classes by the *non-eruptive* classes at $d = 65$. This result demonstrates that although the shallow dynamics of the flank eruptive processes can be different (opening fracture systems at different quotas and with different extension), the presence of deep earthquakes prior to an eruption is a persistent feature. It seems possible that deep seismicity patterns are linked to flank eruption occurrences.

These results are consistent with observations carried out during some months before the main flank eruption of 1991–1993 (Rymer et al., 1993; Murray 1994; Nunnari and Puglisi, 1994; Bonaccorso et al., 1994) that suggested that the volcanic pile was affected by a deformational event before each of the eruptions considered. Some days before the eruption, the inflation underwent a drastic acceleration (Nunnari and Puglisi, 1994; Bonaccorso et al., 1994, 1996) and the magma upraise was accompanied by a rapid failure of the host rocks (Patanè et al., 1996). Unfortunately similar evidence is not available for the other eruptions. However, the results from our MHT and EDT analyses suggest that the occurrence of 'deep' swarms before a flank eruption onset could act with a surprising regularity, i.e. that deep earthquakes occurrence can be considered a 'marker' of the forthcoming flank eruption onset.

(2b) The *eastern flank* earthquake variable has a *T*-ratio less than 1 for all the time points and hence this sub-group is not likely to be linked to eruptive activity. EDT analysis discriminates five

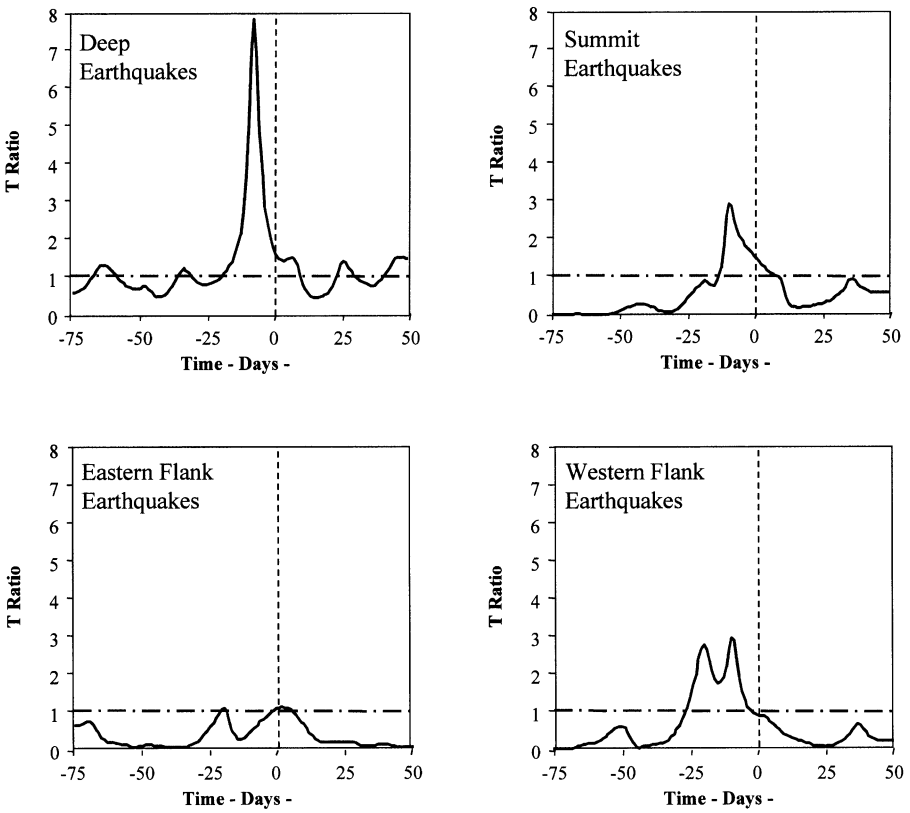

Fig. 7. Temporal evolution of *T*-ratios for the four different seismic data subgroups, considering duplicate 1 as the non-eruptive class. The dashed line indicates the onset of eruptive processes.

eruptions (1981, 1983, 12/1985, 1986 and 1991–1993) at $d = 75$ days, while the remaining three eruptions (1984, 3/1985, 1989) are discriminated at $d = 6$. This lack of distinction reflects the instability of the eastern flank area, the different mechanics found here from those detected elsewhere on the volcano edifice (McGuire et al., 1990, 1997; Lo Giudice and Rasà, 1992; Borgia et al., 1992), and the complex heterogeneous stress field (Cocina et al., 1997). EDT results suggest that the intense seismic swarms occurring at the eruption onset and accompanying the opening of major fracture systems are not sufficient to clearly distinguish the *eruptive* classes from the *non-eruptive* classes. A second discriminating variable is found at the beginning ($d = 6$) of the time samples considered. The instability of EDT results are a further indication that the bulk of seismicity seems to be related to movement on well-known fault systems affecting this area (Montalto et al., 1996; De Rubeis et al., 1997; Gresta et al., 1997; Latora et al., 1999), rather than with the eruptive dynamics of the volcano.

(2c) The *summit* earthquakes are characterized by a marked *T*-ratio greater than 1 in the period from 15 to 0 days before the eruptive episode. EDT analysis discriminates the *eruptive* classes from the *non-eruptive* classes at $d = 70$. These values are similar according to the *T*-ratios computed for the *deep* group and the EDT discriminating variable for this group ($d = 65$). As the events cluster around the conduits in the summit area, accompanying magmas transfer towards the opening fracture systems during the last ascending phases (Castellano et al., 1993). The 5-day time lag between activation of the *deep* and the *summit*

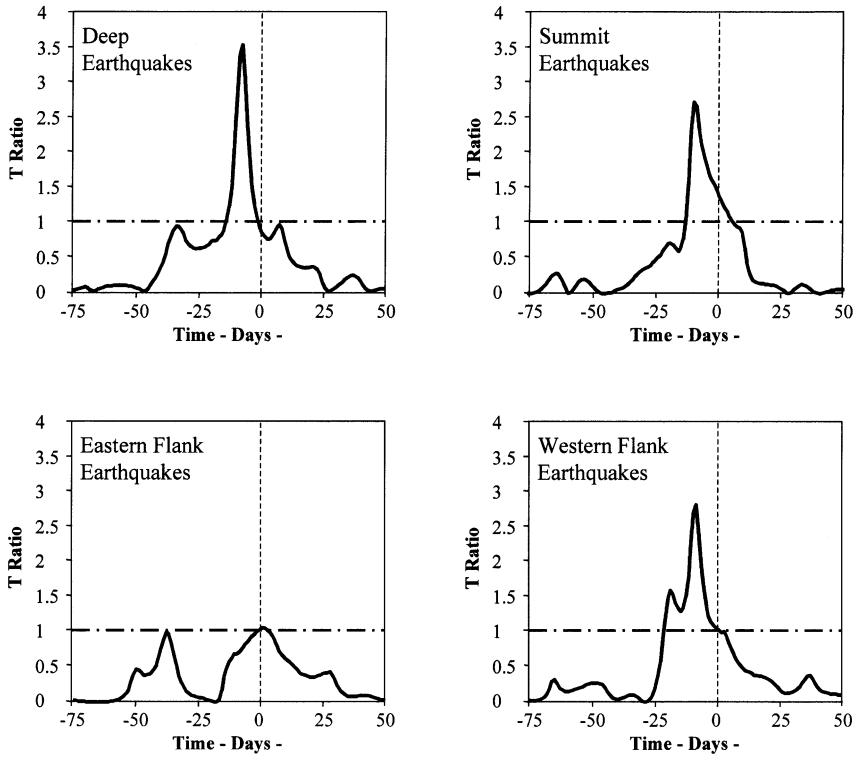

Fig. 8. Same as in Fig. 6, considering duplicate 2 as the non-eruptive class.

earthquakes can be interpreted as due to the time it takes for magma to reach the surface.

(2d) The analysis of the *western* flank earthquakes reveals a very interesting result. The *T*-ratio is greater than 1 from 25 to 5 days before the eruptive episodes and the EDT analysis discriminates well, the two classes at $d = 55$, i.e. 20 days before the eruption onset. Swarm occurrences on the western flank have been observed before some eruptions, inducing some authors to suggest that movement on the NE–SW tectonic system affecting this part of the volcano is related to the volcanic activity (Gresta et al., 1990; Bonaccorso et al., 1996). Magma could rise along intrusions emplaced in the NE–SW fault system (Gresta et al., 1990; Bonaccorso et al., 1996) or reside in magma pockets located in the western flank before the final ascent (Patanè et al., 1996). Our analysis cannot distinguish between these alternatives, but the 25 to 5 day time period found by EDT analysis suggests that the activation of shallow structures in this sector (mainly trending NE–SW) can be considered a 'marker' of volcanic activity inside the time span considered.

6. Conclusions

We have used two different statistical methods, the MHT and EDT to identify and discriminate seismicity patterns affecting the main seismogenic volumes at Mt. Etna volcano prior to flank eruptions during 1981–1996. Though based on different philosophies, they have given consistent results revealing a typical pattern in the 'deep' and 'western' flank earthquakes prior to eruptive activity. In detail:

(1) The analysis of the *total* number of earthquakes identifies an increase in the daily event rate 20 days before flank eruption onsets and discriminates seismic patterns seven days prior to flank eruption onsets. This result provides a constraint, in terms of earthquake occurrence, on the accelerating processes preceding flank eruptions at Etna.

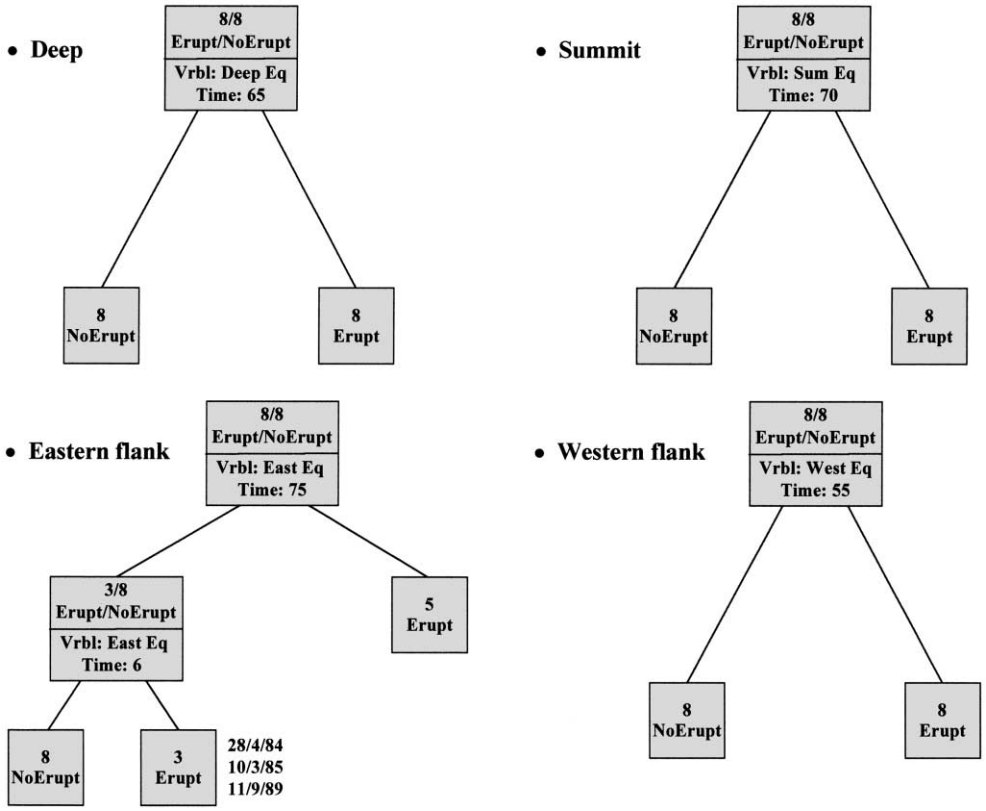

Fig. 9. Entropic decision tree for the four different seismic data subgroups.

(2) *Deep* earthquake analysis identifies an increase in the daily event rates 20 days before flank eruption onsets and discriminates seismic patterns 10 days prior to flank eruption onsets. Deep earthquake occurrence is a marker of volcanic activity and pattern recognition analysis suggests that this process acts with a surprising regularity and could be used as a diagnostic tool for middle to short-term volcano hazard evaluation.

(3) *Eastern flank* earthquake analysis did not identify any significant seismic pattern preceding flank eruption occurrence. Seismic patterns are discriminated the day before the onset for five of the eight eruptions considered in the present study. Earthquakes do not show any significant relation with the eruptive episodes, a result that is consistent with the extreme heterogeneity of the stress field of this seismogenic volume and the different mechanics of the eastern flank area with respect to the whole volcano edifice.

(4) The analysis of the *western* flank earthquakes identifies an increase in the daily event rates 25 to 5 days before flank eruption onsets and discriminates seismic patterns 20 days prior to flank eruption onsets. This seismogenic volume is more homogeneous than the eastern flank volume and the fault systems affecting this sector could play a basic role, acting in a systematic way in relation to the eruptive feeding system.

Acknowledgements

S. Day is acknowledged for revision of the manuscript. Seth Moran and one anonymous referee greatly improved the paper. V.L. thanks CNR and Blanceflor Foundation for financial support. S.V. thanks the Center for Theoretical Physics and the Department of Chemical Engineering of MIT for their kind

hospitality. S.V. has been supported by MURST 60% funds. M.S. Barbano is acknowledged to have encouraged the present study.

References

Armienti, P., Innocenti, F., Petrini, R., Pompilio, M., Villari, L., 1989. Petrology and Sr–Nd isotope geochemistry of recent lavas from Mt. Etna: bearing on the volcano feeding system. J. Volcanol. Geotherm. Res. 39, 315–327.

Azzaro, R., Neri, M., 1992. The 1971–1991 eruptive activity of Mt. Etna (first steps to relational data-base realization). CNR, IIV, Open File Report, 3/92, 10 pp, 21 tabb., 16 tavv.

Barbano, M.S., De Rubeis, V., Tosi, P., Vinciguerra, S., 1999. Clustering properties of Etna seismicity during 1981–1991. J. Seismol. 2 (1–6), 1999.

Bonaccorso, A., Velardita, R., Villari, L., 1994. Ground deformation modelling of geodynamic activity associated with the 1991–1993 Etna eruption. Acta Vulcanol. 4, 87–96.

Bonaccorso, A., Ferrucci, F., Patanè, D., Villari, L., 1996. Fast deformation processes and eruptive activity at Mt. Etna (Italy). J. Geophys. Res. 101, 17467–17480.

Borgia, A., Ferrari, L., Pasquarè, G., 1992. Importance of gravitational spreading in the tectonic and volcanic evolution of Mount Etna. Nature 357, 231–234.

Bousquet, J.C., Lanzafame, G., Paquin, G., 1988. Tectonic stresses and volcanism: in situ stress measurement and neotectonic investigations in the Etna area (Italy). Tectonophysics 149, 219–231.

Cardaci, C., Coviello, M., Lombardo, G., Patanè, G., Scarpa, R., 1993. Seismic tomography of Etna volcano. J. Volcanol. Geotherm. Res. 56, 357–368.

Castellano, M., Ferrucci, F., Godano, C., Imposa, S., Milano, G., 1993. Upwards migration of seismic focii: a forerunner of the 1989 eruption of Mt. Etna. Bull. Volcanol. 55, 357–361.

Castellano, M., Bianco, F., Imposa, S., Milano, G., Menza, S., Vilardo, G., 1997. Recent deep earthquake occurrence at Mt. Etna (Sicily, Italy). Phys. Earth Planet. Inter. 102, 277–289.

Cheung, J., Stephanopoulos, G., 1990. Representation framework. Comput. Chem. Eng. 14, 494–510.

Cocina, O., Neri, G., Privitera, E., Spampinato, S., 1997. Stress tensor computations in the Mount Etna area (Southern Italy) and tectonic implications. J. Geodyn. 23, 109–127.

Cristofolini, R., Gresta, S., Imposa, S., Patanè, G., 1985. Feeding mechanism of eruptive activity at Mt. Etna based on seismological and petrological data. In: King, N., Scarpa, R. (Eds.), Modelling of Volcanic Processes. Earth Evol. Sci., 73–93.

De Rubeis, V., Tosi, P., Vinciguerra, S., 1997. Time patterning properties of the Etna region seismicity during 1874–1913. Geophys. Res. Lett. 24, 2331–2334.

Ferrucci, F., Rasà, R., Gaudiosi, G., Azzaro, R., Imposa, S., 1993. Mt. Etna: a model for the 1989 eruption. J. Volcanol. Geotherm. Res. 56, 35–55.

Frazzetta, G., Villari, L., 1981. The feeding of the eruptive activity of Etna volcano. The regional stress field as a constraint to magma uprising and eruption. Bull. Volcanol. 44, 269–282.

Gasperini, P., Gresta, S., Mulargia, F., 1990. Statistical analysis of seismic and eruptive activities at Mt. Etna during 1978–1987. J. Volcanol. Geotherm. Res. 40, 317–325.

Gresta, S., Patanè, G., 1983a. Variation of *b* values before the Etnean eruption of March 1981. Pure Appl. Geophys. 121, 287–295.

Gresta, S., Patanè, G., 1983b. Changes in *b* values before the Etnean eruption of March–August 1983. Pure Appl. Geophys. 121, 903–912.

Gresta, S., Glot, J.P., Patanè, G., 1985. Studio di meccanismi focali di terremoti etnei. Proceedings of the Fourth Annual Meeting G.N.G.T.S. CNR-Rome, pp. 809–815.

Gresta, S., Patanè, G., 1987. Review of seismological studies at Mount Etna. Pure Appl. Geophys. 125, 951–970.

Gresta, S., Longo, V., Viavattene, A., 1990. Geodynamic behaviour of eastern and western sides of Mount Etna. Tectonophysics 179, 81–92.

Gresta, S., Longo, V., 1994. An attempt at identifying seismological precursor for flank eruptions at Mt. Etna volcano. Acta Vulcanol. 5, 187–191.

Gresta, S., Mulargia, F., Marzocchi, W., 1994. Is there a correlation between larger local earthquakes and the end of eruptions at Mount Etna volcano, Sicily?. Geophys. J. Int. 116, 230–232.

Gresta, S., Bella, D., Musumeci, C., Carveni, P., 1997. Some efforts on active faulting processes (earthquakes and aseismic creep) acting on the eastern flank of Mt. Etna (Sicily). Acta Vulcanol. 9, 101–108.

Guest, J.E., Chester, D.K., Duncan, A.M., 1984. The Valle del Bove, Mount Etna: its origin and relation to the stratigraphy and structure of the volcano. J. Volcanol. Geotherm. Res. 21, 1–23.

Hirn, A., Nercessian, A., Sapin, M., Ferrucci, F., Wittlinger, G., 1991. Seismic heterogeneity of Mt. Etna: structure and activity. Geophys. J. Int. 105, 139–153.

Kamimura, R.T., 1997. Application of multivariate statistics to fermentation database mining. PhD thesis, Massachusetts Institute of Technology.

Kieffer, G., 1985. Evolution structurale et dinamique d'un grand volcan poligenique: stades d'édification et activité actuelle de l'Etna (Sicile). PhD thesis, University of Clermont Ferrand, 497 pp.

Klein, F.W., Koyanagi, R.Y., Nakata, J.S., Tanigawa, W.R., 1987. The seismicity of Kilauea's magma system. U.S. Geol. Surv. Prof. Pap. 1350, 1019–1186.

Latora, V., Rapisarda, A., Vinciguerra, S., 1998. A fractal approach to the temporal distribution of microseismicity at the low eastern flank of Mt. Etna during 1989–1994. Phys. Earth Planet. Inter. 109, 115–127.

Latora, V., Vinciguerra, S., Bicciato, S., Kamimura, R.T., 1999. Identifying seismicity patterns leading flank eruptions at Mt. Etna Volcano during 1981–1996. Geophys. Res. Lett. 26, 2105–2108.

Lee, W.H.K., Lahr, J.C., HYPO71 (revised), 1975. A computer program for determining hypocenter, magnitude, and first motion pattern of local earthquakes. U.S.Geol.Surv., Open-file Report, pp. 75–311.

Lo Giudice, E., Patanè, G., Rasà, R., Romano, R., 1982. The structural framework of Mt. Etna. In: Romano, R. (Ed.), Mt. Etna Volcano. Mem. Soc. Geol. Ital. 23, 154–176.

Lo Giudice, E., Rasà, R., 1992. Very shallow earthquakes and brittle deformation in active volcanic areas: the Etnean region as example. Tectonophysics 202, 257–268.

Mardia, K.V., Kent, J.T., Bibby, J.M., 1979. Multivariate Analysis. Academic Press, San Diego.

McGuire, W.J., Pullen, A.D., Saunders, S.J., 1990. Recent dyke-induced large-scale block movement at Mount Etna and potential slope failure. Nature 343, 357–363.

McGuire, W.J., Moss, J.L., Saunders, S., Stewart, I.S., 1997. Intravolcanic rifting at Mount Etna in the context of regional tectonics. Acta Vulcanol. 9 (1/2), 147–156.

Montalto, A., Vinciguerra, S., Menza, S., Patanè, G., 1996. Recent seismicity of Mount Etna: implications for flank instability. In: McGuire, W.J., Jones, A.P., Neuberg, J. (Eds.), Volcano Instability on the Earth and Other Planets. Geological Society Special Publication 110, pp. 169–177.

Mulargia, F., Gasperini, P., Marzocchi, W., 1991. Pattern recognition applied to volcanic activity: identification of the precursory patterns to Etna recent flank eruptions and periods of rest. J. Volcanol. Geotherm. Res. 45, 187–196.

Mulargia, F., Marzocchi, W., Gasperini, P., 1992. Statistical identification of physical patterns to Etna recent flank eruptions and periods of rest. J. Volcanol. Geotherm. Res. 53, 289–296.

Murray, J.B., 1994. Elastic model of the actively intruded dyke feeding the 1991–93 eruption of Mt. Etna derived from ground deformation measurements. Acta Vulcanol. 4, 97–99.

Nercessian, A., Hirn, A., Sapin, A., 1991. A correlation between earthquakes and eruptive phases at Mt. Etna: an example and past occurrences. Geophys. J. Int. 105, 131–138.

Nunnari, G., Puglisi, G., 1994. The global positioning system as useful technique for measuring ground deformations in volcanic areas. J. Volcanol. Geotherm. Res. 61, 267–280.

Patanè, G., Gresta, S., Imposa, S., 1984. Seismic activity preceding the 1983 eruption of Mt. Etna. Bull. Vulcanol. 47, 941–952.

Patanè, D., Caltabiano, T., Del Pezzo, E., Gresta, S., 1992. Time variation of *b* and Qc coefficients at Etna volcano (1981–1987). Phys. Earth Planet. Inter. 71, 137–140.

Patanè, G., Montalto, A., Imposa, S., Menza, S., 1994. The role of regional tectonics, magma pressure and gravitational spreading in earthquakes of the eastern sector of Mt. Etna volcano (Italy). J. Volcanol. Geotherm. Res. 61, 253–266.

Patanè, G., Montalto, A., Vinciguerra, S., Tanguy, J.C., 1996. A model of the 1991–1993 eruption onset of Etna. Phys. Earth Planet. Inter. 97, 231–245.

Quinlan, J.R., 1986. Induction of decision tree. Mach. Learning 1, 81–106.

Rasà, R., Azzaro, R., Leonardi, O., 1996. Aseismic creep in faults and flank instability at Mount Etna volcano, Sicily. In: McGuire, W.J., Jones, A.P., Neuberg, J. (Eds.), "Volcano Instability on the Earth and Other Planets". Geological Society Special Publication 110, pp. 179–192.

Rymer, H., Murray, J.B., Brown, G.C., Ferrucci, F., McGuire, W.J., 1993. Mechanisms of magma eruption and emplacement of Mt. Etna between 1989 and 1992. Nature 361, 439–441.

Scarpa, R., Patanè, G., Lombardo, G., 1983. Space–time evolution of seismic activity at Mount Etna during 1974–1982. Ann. Geophys. 1, 451–462.

Sharp, A.D.L., Lombardo, G., Davis, P.M., 1981. Correlation between eruptions of Mount Etna, Sicily, and regional earthquakes as seen in historical records from AD 1582. Geophys. J. R. Astron. Soc. 65, 507–523.

Shannon, C., 1949. The Mathematical Theory of Communication. University of Illinois Press, Urbana.

Shlien, S., Toksoz, M.N., 1970. A clustering model for earthquake occurrence. Bull. Seismol. Soc. Am 60, 1765–1787.

Smithsonian Institution, 1981, 1984, 1985, 1986, 1989, 1991. SEAN Bullettin. Smithsonian Institution, Washington DC.

Tinti, S., Mulargia, F., 1985. An improved method for the analysis of the completeness of a seismic catalogue. Lett. Nuovo Cimento 42, 21–27.

Vinciguerra, S., Gresta, S., Barbano, M.S., Distefano, G., 1999. A comparison between the *b* value and the fractal dimension of the Mt. Etna seismicity during 1983–1996, submitted for publication.